

DOCUMENT RESUME

ED 294 409

EC 202 580

AUTHOR Jones, Eric D.; Southern, W. Thomas
TITLE Chronological Age at School Entrance and the Prevention of Learning Disabilities: Policy Making and the Misinterpretation of Research.
PUB DATE Sep 87
NOTE 30p.; Paper presented at the Annual Meeting of the Council for Learning Disabilities (San Diego, CA, September 1987).
PUB TYPE Speeches/Conference Papers (150) -- Information Analyses (070)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Age Grade Placement; *Early Admission; Educational Needs; Gifted; Grade Repetition; *Learning Disabilities; Primary Education; *Research Methodology; *School Entrance Age; *School Readiness; Student Placement

ABSTRACT

This paper (1) presents a critical analysis of the literature on school readiness and school failure; (2) discusses the problems that ill conceived policies based upon readiness studies present to the education of learning disabled, regular class, and gifted children; and (3) offers guidelines for policies and programs. Methodological issues raised by the analysis of school readiness research include experimenter bias, errors in sampling procedures or sampling description, matching errors, sample restriction, misinterpretation of identified statistical relationships, and selection of dependent variables. Misinterpretation of the studies has resulted in magnification of the importance of obtained differences, confusing correlation and causation, concluding that no treatment (keeping the student out of school another year) is preferable to early intervention, and misapplying the findings to early entrance or acceleration of gifted students. It is concluded that arbitrary cutoff birthdates as a basis for determining school entry should be replaced by assessing the needs and characteristics of all children and providing them with an individually appropriate educational program. (Author/DB)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

**Chronological age at school entrance
and the prevention of learning disabilities:
Policy making and the misinterpretation of research**

Eric D. Jones and W. Thomas Southern
Bowling Green State University

Paper presented at the annual meeting of the
Council for Learning Disabilities, San Diego, 1987

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Eric D. Jones
Jones

TO THE EDUCATIONAL RESOURCES

Abstract

Studies of school readiness and school failure are over-interpreted and fraught with methodological problems. This paper (a) presents a critical analysis of the literature on school readiness and school failure, (b) discusses the problems that ill conceived policies based upon readiness studies present to the education of LD, regular class, and gifted children, and (c) offers guidelines for policies and programs.

**Chronological age at school entrance and
the prevention of learning disabilities: Policy making and
the misinterpretation of research**

The best age for a child to begin school has been an issue that has concerned practitioners, policy makers, and perhaps more sporadically educational researchers. The discussion of the issue is also a curiosity for educational researchers interested in the influence of research on educational policy and practices. Two conceptually related bodies of research on the importance of chronological age have accumulated separately and virtually without agreement. Studies of school readiness have generally revealed discouraging prospects for the academic success of children who enter school at a relatively early age. The results of those studies have prompted some researchers and writers (e.g. Diamond, 1983; DiPasquale et al, 1980; Uphoff & Gilmore, 1985, 1986a, 1986b), to take the position that the immaturity of early entrants puts them at substantially greater risks for a host of calamities including: academic underachievement, grade retention, and even referral to programs for learning disabled students. On the other hand, studies of the effects of promoting gifted young children have frequently found positive effects for acceleration. None of the acceleration studies found negative effects for academically capable students. It is unusual to find two bodies of research that are so closely related in their implications for educational policies and practices, but so contradictory in their results.

The purpose of this paper is to evaluate the literature on the effects of early entrance on achievement and social-psychological adjustment. The studies from the school-readiness and gifted education will be reviewed and summarized. Implications of the studies for policy making in the areas of prevention of learning disabilities and nurturing of giftedness will be identified and

discussed.

School Readiness Research

Studies of academic success and social adjustment of children who are relatively young compared to their classmates report that younger children are more likely to (a) reach lower levels of academic achievement (Baer, 1958; Bigelow, 1934; Carroll, 1963; Carter, 1956; DiPasquale, Mule, & Flewelling, 1980; Forester, 1955; Hall, 1963; King, 1955; Langer, Kalk, & Searls, 1984; Uphoff & Gilmore, 1985, 1986a, 1986b), (b) have more difficulties in social and emotional adjustment (Baer, 1958; DiPasquale et al, 1980; King, 1941), and (c) be at greater risk for referral for retention (DiPasquale et al, 1980; Langer et al, 1984) or placement in programs for children with learning disabilities (Diamond, 1983; DiPasquale et al, 1980). Several writers (e.g. Donofrio, 1977; Frick, 1986; Uphoff & Gilmore, 1985, 1986a, 1986b) have argued: (a) against early school admission of young children regardless of their apparent aptitude, (b) in favor of raising the minimum age for school entrance, and (c) in favor of recommending that children who appear to be immature at preschool screenings be held out of school an additional year. These arguments are based upon the grounds that even when young children are above average in cognitive ability they are not able to overcome the disadvantage of being younger and, therefore, they are less ready for school compared to their relatively older classmates.

The validity of conclusions that elementary school children who are younger than their classmates are at greater risk of academic failure or social-psychological difficulties are dependent upon the quality of the research methodologies that have been used. Our review of the studies identified many serious methodological errors -- enough to lead us to consider that the conclusions drawn by critics of early entrance are virtually without foundation.

Methodological Issues

The studies from the school readiness literature are in most cases based upon seriously flawed

methodologies. The methodological flaws pertain to sampling, procedural bias, statistical conclusion validity, and the selection of dependent variables. The limitations that each of the flaws place upon the implications drawn from these studies will be discussed.

Experimenter bias One of the most important threat to the validity of a study is the degree to which experimenter bias could have influenced the sample selection, data collection or reporting of the results. It is quite possible that several of the school readiness studies, as well as the fugitive studies cited in the reports to bolster the conclusions of the authors, were unduly influenced by political motivations. Each of the between-group studies that observed negative effects of early entrance (Baer, 1958; Bigelow, 1934; Carroll, 1963; Carter, 1956; Forester, 1955; Hall, 1963; King, 1955; Uphoff & Gilmore, 1985, 1986a, 1986b) were conducted by persons in the employ of the school districts in which the studies were conducted. The introductions of several reports (Baer, 1958; Bigelow, 1934; Carroll, 1963; Carter, 1956; Forester, 1955; Hall, 1963; King, 1955) referred to local controversies over how old children should be when they enter school. It appears that in each district administrators and teachers were opposed to early entrance, but were confronted by parental pressures to admit some students before they met the minimum age for school entrance. Of the five studies that observed no negative effects as a result of early entrance, four studies (Deitz & Wilson, 1985; Green & Simmons, 1967; May & Welch, 1984; Miller & Norris, 1967) were conducted by researchers who were employed outside the schools in which the studies were conducted. Only one study that observed no ill effects due to early entrance (Partington, 1937) was conducted by someone already employed by the local school district. Further hints of bias are revealed in the introductions of the reports. For example, in the studies by independent researchers (Deitz & Wilson, 1985; May & Welch, 1984; Miller & Norris, 1967) and in the study by Partington (1937) the development of the arguments in favor of and against early entrance were more even-handed than in the school district studies.

Sampling. Two general procedures have been used most frequently to assemble samples for study. Both procedures contributed to threats to the internal validity of the studies of school readiness. We do not intend to suggest that all or even most of the district level studies that reported negative effects due to early entrance were politically oriented pseudo-evaluations (cf. Stufflebeam & Webster, 1980), but it is likely that some of them were. Southern and Jones (1987) found that elementary school teachers and principals are generally opposed to early entrance or acceleration for young gifted children. It is not implausible for the prevailing biases of school personnel to influence the results of local evaluations.

The first and most commonly used procedure has been to compare the performances of a sample of children who entered school before the prevailing entrance age for the district with the performance of a group of students who were not underage for school entry (Baer, 1958; Bigelow, 1934; Carroll, 1963; Carter, 1956; Forester, 1955; Green & Simmons, 1967; Hall, 1963; King, 1955; Uphoff & Gilmore, 1985, 1986a, 1986b). All of these studies dealt with nonrandom samples of convenience from individual districts.

Sampling procedures. Three threats to the validity of the studies can be attributed to procedures used in the selection of samples including: (a) inadequate description of samples, (b) inappropriate matching of independent variables, and (c) restriction of samples.

Sample description. Most studies provided only the barest description of the characteristics of the samples. Generally, the only independent variables that were described were: grade level, age, sex, and intelligence. Many of the studies dealt with small samples, and understandably could not measure the effects of many variables, but it would have been possible for the authors to have provided substantially more informative descriptions of their samples. Without adequate descriptions of samples it is virtually impossible to assess the degree to which generalizations can be made from the results of the studies.

Matching. Matching procedures were frequently used to form comparison groups of older students to compare with the early entrants at each grade level. Experimenter bias can of course influence the selection of matched samples, but there are more mundane problems associated with the use of matching procedures. The purpose of matching is to increase the similarity between groups on independent variables that are known to be (or presumed to be) related to performance on the dependent measure. There are problems with matching, because it is rare that one has sufficient knowledge before the study is conducted to be certain that a given variable will have a demonstrable effect on the dependent variable. Therefore, most cases matching are based upon presumptions of relationships between selected independent variables and the dependent variables. It is critically important that the presumptions of relationships be critically evaluated, because it is quite possible that the dependent variable will covary with some yet uncontrolled variable.

A clear illustration of the problem of matching without identifying the appropriate variable upon which to match is presented in the school-readiness literature. It appears that in the case of school readiness studies that used a priori matching procedures (Baer, 1958; Carter, 1956; Green & Simmons, 1962; King, 1955), or determined group equivalence with ex post facto analyses (Diamond, 1983; DiPasquale et al, 1980; Forester, 1955; Hall, 1963; Langer, et al, 1984; Miller & Norris, 1967), most of the researchers did not critically evaluate their assumptions of the relationship matching variables and measures of the effects on early entrance. Some studies matched the comparison group to the group of early entrants on the bases of IQ and sex (Baer, 1958; Carter, 1956; Green & Simmons, 1962). Instead of matching several other studies (Forester, 1955; King, 1955; Partington, 1937, Miller & Norris, 1967) reported data on the comparability between the groups on IQ and sex.

In the introduction to his study, Carter (1956) identified several reasons that have been used to justify the early admission of children in the Austin Public Schools. Three of the reasons

(employment of mothers during the day, inadequate living conditions at home, and inadequate play facilities in the neighborhood) are related to socio-economic status. Children from lower socio-economic status families are more apt to come from inadequate home situations, neighborhoods that lacked adequate play facilities, and have mothers who worked outside the home.

Socio-economic status has been widely observed to have a persistent influence on achievement and school adjustment. Studies that ignore the effects of socio-economic status run the risk of finding that young children from lower socio-economic status families who enter school early tend to have lower levels of academic achievement than older middle class children who do not enter school later.

Unfortunately, socio-economic status has been ignored in most studies of the effects of early entrance. Only two studies attempted to measure or control the effect of socio-economic status. Green and Simmons (1962) matched early entrants and their older classmates on the basis of socio-economic status. They observed that there was no verifiable advantage to barring capable young students from early school entrance. Langer et al (1984) examined the influence of socio-economic status on the achievement of early entrants. The study by Langer et al (1984) observed a statistically significant ($r = -.055$, $p < .01$; $n = 114,000$), but trivial correlation between achievement and school entrance age, but their data also revealed that the correlations between achievement and: (a) the quality of the home environment, (b) level of parental education, and (c) the quality of the community accounted for a larger portion of the variance in achievement test scores than school entrance age. The results of the studies by Green and Simmons (1962) and Langer et al (1984) indicate that socio-economic status may be an important independent variable in the research on the effects of early entrance and should be considered in either the description or equating of samples.

 Insert Tables I and II from Langer et al (1984) about here

Sample restriction. The samples used in the school readiness studies have been formed by identifying the children in different grade levels who were (a) given early entrance, (b) born during the summer months, or (c) were ranked from youngest to oldest. None of the studies systematically investigated the effects of early entrance on children who were selected for early entrance because they were academically accelerated. In fact, the frequently cited studies by King (1955) and DiPasquale et al (1980) purposefully eliminated from their studies children who were significantly above average intelligence as well as children in special education programs. On the other hand, the results of studies based upon restricted samples have been freely generalized to the population of gifted young children.

Statistical conclusion validity. Tests of statistical significance are frequently used to determine whether or not relationships between variables are sufficiently strong that one could reasonably conclude that chance alone would not account for the relationships. Statistical tests do not prove that relationships actually exist they merely indicate the degree to which chance factors alone would be expected to account for the relationships. If the probability that chance factors would explain the relationship is low, then the researcher generally concludes that the relationship between the variables is systematic. The validity of conclusions based upon statistical analyses are limited to the extent to which randomly occurring relationships do not appear to be systematic. Tom Robbins (1984) description of a pagan cemetery illustrates the difficulty in determining which relationships occur randomly and which are systematic.

The burial mound was outside the city walls in a field dotted with cow pies and large stones. The stones had been arranged geometrically in patterns that were supposed to

to mean something to the gods. Presumably, the cow pies had fallen at random, although then, as now, the division between what is random in nature and what is purposeful is extremely difficult to determine. (*Jitterbug Perfume*, pp. 29-30)

In order to proceed with power and confidence in our understandings of events, it critically important to determine the extent to which events can reasonably be expected to occur in predictable relationships to one another. The difference between random and nonrandom events has been blurred by the school-readiness researchers who (to persist in Robbins image) have attempted to find patterns and meanings in the arrangement of cow pies. Cook and Campbell (1979) identified three factors that threaten the validity of conclusions that are based upon the results of statistical analyses. Two of those factors - "fishing rate" and sample size - can be found in the school readiness evaluations of the effects of early entrance.

The first threat to the statistical conclusion validity in school-readiness studies is the referred to as the fishing rate problem. A basic rule of scientific investigation is that hypotheses be stated first at the outset of the study. It is not proper to gather a set of data and proceed to run statistical analyses until interesting relationships are discovered, and then report the relationships as though they were initially hypothesized. The reason hypotheses must be stated a priori is because, the more the investigator "fishes" for relationships the greater the probability that randomly occurring relationships will appear, by chance, to be systematic differences. The probability that a coincidental relationship will appear to be systematic can easily be calculated. For the t -tests and one-way analyses of variance that were frequently used in these studies, simply multiply the level selected for determining statistical significance (usually $\alpha = .05$) by the number of analyzed that were conducted. For example, in the case of the study by Baer (1958) it appears that approximately 70 simple analyses of variance were conducted. It is a virtual certainty that Baer (1958) would have found some statistically significant relationships by chance alone. Other studies that made slightly less prodigious pursuits of statistically

significant relationships included: (Bigelow, 1934; Carroll, 1963; Carter, 1956; Diamond, 1983; DiPasquale et al, 1960; King, 1955; Miller & Norris, 1967).

The study by Uphoff and Gilmore (1985; reprinted 1986a, 1986b) is open to an interesting, if unsophisticated, variation on the "fishing rate" error. Uphoff and Gilmore made many comparisons between children who were born in the months of summer and early fall and their classmates who were born earlier in the year. Uphoff and Gilmore (1985) reported several differences in the relative performances of the two groups. They did not, however, carry their analyses beyond simple intuitive comparisons of frequencies. Large as some differences in frequencies may appear to be, the differences may not be greater than would occur by chance alone. Miller and Norris (1967), for example, compared the academic achievements of children who were relatively young for their grade with the performances of their relatively older classmates. The frequency data reported by Miller and Norris (1967) appear at first glance to be large enough to suggest some systematic differences between the groups. In several of the analyses, however, Chi-square analyses failed to reveal statistically significant differences. It is likely that the differences that Uphoff and Gilmore (1985) alleged would not be statistically significant either. It is incumbent upon researchers to analyze their data if they are going to claim statistical significance. Given that Uphoff and Gilmore (1985) chose not to analyze their data, it has to be assumed that no significant relationships existed. Biased and impressionistic analyses of sets of raw data contribute substantially to the identification of relationships that are in all likelihood due to chance alone.

Sample size. There are several ways that sample size can contribute to threats to the statistical conclusion validity of a study. First, samples that are too small frequently do not allow sufficient power to observe a statistically significant difference, even though the difference is systematic and not random. That problem was not apparent in this collection of published studies. Several studies did use small samples and may have been affected by the second threat. The second problem in dealing

with a small samples is that in small samples a few extreme scores can have substantial effects on the results of analyses. Since the samples for these studies were rarely chosen with the same care that candidates for early entrance are chosen and reviewed it is entirely possible that this problem plagues several of the studies. Unfortunately, only a few studies reported tables of descriptive statistics that would allow for estimates to be made of the within groups variances. Very large samples present a third threat to the claim that systematic relationships actually exist. Two studies (Diamond, 1983; Langer et al, 1984) studied samples in excess of 100,000 subjects. With such large samples it is likely that even the most trivially small systematic difference may be detected. Unfortunately for Diamond (1983) and Langer et al (1984) their studies were so confounded that the differences that they observed could not be clearly attributed to the relative ages of the students.

Selection of dependent variables. The validity of conclusions that relative age-in-grade is somehow systematically related to school performance rests upon the dependent variables upon which performance is judged. Evaluations of the degrees to which young children successfully meet the demands of schooling have been based upon a variety of variables. Most of the school-readiness studies used measures of achievement on standardized achievement tests as the basic criterion to evaluate the effects of early entrance. Forester (1955) examined grades assigned by teachers. Rates of retention were examined in other studies (Hall, 1963; Langer et al, 1984; Miller & Norris, 1967). DiPasquale et al (1980) used rates of referral for special education services measures of academic achievement. An assortment of indicators of social-emotional adjustment have been used to evaluate the effects of early entrance -- leadership (Mawhinney, 1964), social maturity (Forester, 1955; Baer, 1958). The degree to which meaningful evaluations of the relationship between relative chronological age and achievement or social adjustment can be made depends upon the nature of the variables being measured. Each of the dependent variables listed above have serious limitations on their value as

indicators of the success of children who enter school at a relatively early age.

The measures of achievement that have been used are inadequate as the bases for evaluation. Care should also be taken in the interpretations of standardized achievement test results. Gredler (1978) pointed out a basic problem with the use of **norm referenced evaluation** of achievement. Unless the tests have been carefully selected so that curriculum content would be reflected in the content of the standardized test, the content validity of the measure may well be seriously limited. If the content of the test differs markedly from the content of the instructional curriculum, test results would, in the primary grades be apt to provide serious underestimates of the achievement of relatively young students who enter school with less proficiency in academic skills than older children. The task of selecting a test that adequately measure achievement in local curricula is not a simple task. The 1980 annual reports of the Regional Title I (currently Chapter I) Technical Assistance Centers indicate that assistance in district level test selection constituted a substantial portion of their field services.

While standardized achievement tests may vary in the degree to which they are adequate measures of the achievement of local curricula, they are the least biased measures of achievement that have been used to evaluate the effects of early entrance on academic achievement. Teacher assigned grades are apt to be biased. Data on the rates of retention and referral for special education services would perpetuate the biases associated with teacher grades and perhaps reflect additional teacher bias. Gredler (1978, 1980) argued that teachers tend to form norm-referenced types of expectation to which the children in their classrooms must to some degree conform. Similarly Gerber and Semmel (1984) contended that teachers form perceptions about the "teachability" of students in their classrooms. Students who are regarded as being difficult to teach are at greater risk for referral for special education services. Immature students are apt to be regarded by their teachers as being less teachable compared to their more mature peers. Even the most casual observer is aware that the immaturity of an individual is both debatable, and not perfectly correlated with age. The labeling of young children as

immature is, however, less apt to be questioned, and more apt to result in their being considered to be "difficult to teach," thus increasing the probability that they will be candidates for retention or special education services. Children who are relatively young are more vulnerable for referral for retention or for special education services than their older classmates.

Discrepancies between the levels children are expected to achieve and observed levels of achievement are occasionally used to evaluate the wisdom of early entrance placements. In some studies it appeared that early entrants scored well below their expected levels for achievement (Huff, 1984, cited in Uphoff & Gilmore, 1985). On the surface it appears that failure of early entrants to achieve to their capacities is a serious problem with the procedure. There are, however, three problems that limit the validity of discrepancy measures as indicators of the value of early entrance practices. First, studies that calculated the discrepancy between the potential and actual achievement (Green & Simmons, 1962; Partington, 1937; Huff, 1984, cited in Uphoff & Gilmore, 1985) ignore the fact that the validity of discrepancy measures is compromised by the amounts of error associated with each measure -- plus an unknown amount of error associated with the comparison of the two measures. Even when discrepancy scores are based upon the results of the most reliable achievement and aptitude tests, the amount of error that must be considered is too large to allow for the conclusions as to whether or not early entrants have achieved at levels commensurate with their presumed potential. Discrepancies based upon teacher assigned grades, retention rates, or referrals to special education are too apt to have been influenced by the biases of the teachers to be taken seriously. Studies which do not provide any comparison to older classmates (Huff, 1984; cited in Uphoff & Gilmore, 1985) are obviously inadequate, because no standard exists to judge whether too many early entrants have been achieved below their potential. In summary, analyses of the discrepancies between expected and observed achievement can not be adequately done. Even if the problem of measurement error could be overcome, the studies that examined the discrepancies have failed to present convincing evidence that

early entrants fall short of expectations more often than: (a) they would if their entrance had been delayed, or (b) the general population. However unfortunate it may be, unrealized potential is so common that it is practically a proverb.

The **social-emotional adjustment** of early entrants and relatively young children has been examined with little apparent regard for the theoretical significance of the traits chosen for evaluation, the natures of the traits being studied, or the technical quality of the instruments being studied. The traits that have been measured include: social-emotional maturity (Baer, 1958; Forester, 1955; cited in Uphoff & Gilmore, 1985) and leadership (Mawhinney, 1964 cited in Uphoff & Gilmore, 1986). The measurement of those nonacademic/noncognitive traits has been with only a few exceptions impressionistic in the operationalization, and casual in the execution.

Social-emotional maturity has generally been evaluated by comparing the ratings that teachers give the younger children in the class with the ratings that they give the older children. Most studies failed to provide any detail on the instrumentation that was used to make the ratings. An exception was the study by Miller and Norris (1967) which evaluated social maturity through peer ratings. They did not find statistically significant differences for ratings on seven of the nine sociometric variables. The group means differed for the variables "happiness" and "fun-loving". The differences were that children who were entered school late were considered by their peers to be sadder and more serious than their peers who either entered school early or at the normal time. Despite the difficulties of sociometric peer ratings, the procedure is less influenced by teacher bias than teacher ratings - particularly when ratings are conducted in the upper primary and secondary grades.

Teacher ratings of maturity have been used in several studies (Baer, 1958; Forester, 1955). The use of teacher ratings are inadequate measures for two reasons. First, they are highly subject to bias. Teachers are apt to assign the label "immature" to a child who is difficult to teach and also happens to be younger than most of the other children in the class. There are frequently a number of reasons

other than immaturity that would provide more accurate and instructionally defensible explanations for a child's difficulties than the nebulous label immaturity. The second reason why teacher ratings of social-emotional maturity are inadequate measures is that by asking a primary school teacher to rate the maturity of the students in a given class it is almost assured that some children in the class will be judged as being immature. Whether teachers are asked to rank order the students or rate their maturity with a Likert type of scale, they will tend to make their ratings of any given student in reference to the other students in the classroom. The greater tendency for younger children to be labeled as immature is not unexpected and may well be irrelevant to determining the advisability of allowing a child to enter school early. Ratings of immaturity are of importance only to the extent to which they can explain academic failure. Since teacher ratings of maturity are likely to be so thoroughly confounded by other explanations for failure to learn and adjust that they can hardly be relied upon to provide information that is of any value at all.

Teachers have occasionally been asked to rate the **leadership** skills of early entrants (Mawhinney, 1964 cited in Uphoff & Gilmore, 1985). Leadership is perhaps an easier variable to operationalize, but it is unfortunately irrelevant to the evaluation evaluating the effects of early entrance. The point is clearly illustrated in the observations by Forester (1955) and Mawhinney (1964, cited in Uphoff & Gilmore), that only a small proportion of the early entrants were judged to be outstanding leaders. Forester (1955) did not provide any figures, but Mawhinney (1964) observed that only 20 percent of the early entrants were outstanding leaders. These observations can hardly be taken as serious criticisms of early entrance practices for three reasons. First, the definition of leadership suggests that most people are not outstanding leaders. Second, if a child is offered early entrance because of academic achievements there is no reason to consider that he or she should become a leader. Third, considering that: (a) most people are, by definition, not outstanding leaders, and (b) leadership and talent are imperfectly correlated, Mawhinney's (1964) observation that 20 percent of the

early entrants were outstanding leaders indicates that either the early entrants in his study were exceptionally charismatic bunch of kids, or that Mawhinney's criteria for judging outstanding leadership were too liberal.

Building Mountainous Conclusions Out of Data Molehills

The grave methodological and statistical errors associated with the studies reviewed above should urge caution in discussing and drawing conclusions from the study. Yet the researchers reporting these results have not been temperate in their conclusions. Some have been willing to develop causal relationships between birthdate and later school difficulties. Diamond (1984) posits that positive correlations between birth month and later diagnoses of specific learning disabilities can be ascribed to higher temperatures in the months the fetus is carried to term. She cites correlational data of mean monthly temperatures at the Honolulu airport between January (72 F) and August (84 F) and SLD placements as evidence that this is true. If the high temperature is the culprit here this has grave implications of educational difficulties in equatorial zones. If on the other hand, the difference of mean temperatures is at issue, Midwestern states with monthly mean temperature differences exceeding 50-60 degrees should look for extraordinary numbers of SLD placements.

Uphoff and Gilmore (1986) examined the association between age and youth suicide concluding that the increase in suicides in the past 20 years roughly coincides with the "sifting down of the curriculum" apparent in the early grades. They conclude that the demands of the curriculum can cause stress and maladaptation in "too young" children. They make these assertions despite the fact that there has not been a large consistent rise in youth suicides in the last 20 years (Department of Health and Human Services, 1985), and in the face of a stream of national reports that assert that the curriculum has been diluted and devalued in precisely the same period. Interestingly enough, the suicide rate for Sweden is much higher than in the United States among all age groups, yet Swedish children start school at 7 years old.

The prescriptions of these researchers generally fall into four categories. a) raise the school entry age requirements; b) hold out students who are screened or adjudged too immature; c) institute differential entrance dates for males or others more at risk for academic difficulties; d) strongly discourage early admission for able students.

The bases for each of these recommendations rest on the obtained differential performance of younger versus older students in grade, and it is instructive to examine them in light of their internal logic and the justifiability of their conclusions. Solely for the sake of argument, one can temporarily suspend objections to the accuracy and applicability of the data in the studies. Even when this is done, each of the studies, to a greater or a lesser degree, derives its conclusions under errors that render the results unwarranted, and at bottom absurd.

Magnification of the Importance of the Obtained Differences

While many of the studies reviewed were of small, limited samples, and made no attempt to ascertain other reasons for school failure, a few exist that look at large samples and attempt to identify other factors associated with lower achievement, retention, and referral to special education. Langer et. al. (1984) analyzed a national sample of 114,000 students, looking at a number of variables such as average age in the grade, parent income, parent education, etc. The study looked separately at blacks and whites, and the tables below report the results of regression analysis of each racial group.

Insert Tables I and II from Langer et al (1984) about here

While age was determined to significantly contribute to the prediction of achievement for both populations it is interesting to note the relative size of that effect when compared to variables such as parent education, home environment, or even the region where the student attends school. The relatively small effect of age adds minimally to the prediction of school success. Far more important

are variables directly related to socio-economic status. Indeed, the study cited here treats age as if it were independent of SES and this is probably an unwarranted assumption.

As indicated earlier, parents from lower SES groups have strong reasons for enrolling their children as early as possible. Their reasons include: single parent work, or nonexistent, inadequate, or expensive day care. Consideration of these two variables as correlative may lower still further the significance of the age effect. A further reduction in the significance of the findings might occur if the researchers had elected to examine blacks and whites in a single sample, using race as a predictor in the regression. Given the striking difference in overall retention rates between the samples (16.54 % for whites and 26.16% for blacks at age 13), much of the prediction for achievement and retention might be predicted by the race of the student. Since black males in early grades are retained much more than whites or black females, a strong sex-by-race interaction might further attenuate the importance of the age effect.

All things being equal, one would expect that younger children in a grade might perform less well than older children given that they have a slightly younger mental age on the average. One would also expect this advantage to lower in importance as students progressed through the grades because the ratios of the differences between the chronological age of older and younger children would decrease over time. This is precisely what Langer et. al (1984) report. For example, they found that for the 13 year old samples, class age effects had disappeared. For 17 year olds, "neither age nor class age was significant in the presence of the other predictors". The results of the Langer et. al. (1984) study have been echoed in other studies. Baer, (1958) reports that the achievement of younger students is steady and at the average level for their grade throughout the period he examined. Given the small nature of the effect observed and the transitory nature of much of the differences, it seems unwarranted to prescribe the large scale changes in policy described in the last section.

Correlation and Causation

Some of the researchers seem to have fallen into the logical fallacy of ascribing causality to correlational results. Diamond (1983), Uphoff and Gilmore (1985), Ilg and Ames (1965) etc. certainly have constructed cases dependent on a causal link between age and negative academic and social effects for children. Some of these assumptions lack even a superficial causational rationale. For example, Uphoff asserts that immaturity may be related to the higher levels of off task behavior that he noted among students in his study. He asserts that this behavior causes interference with the learning task. Interestingly, he noted that very young girls deviated more from older students in the class in this behavior than did boys. This seems at odds with his earlier conclusions that early age admission is a more severe problem for boys than for girls.

Most authors in these studies do not bother to define the direct link between age and various outcomes. (Diamond's 1983 article linking birth date and weather to increasing likelihood of learning disabilities is an exception). For these researchers, the relationship established through the correlation is powerful enough to stand as a reason. Such causal constructions are not warranted if the results can be explained in other equally plausible ways.

Another Explanation: Much of the effect observed may be related to differences in the way teachers view children entering their early grades. Teachers may have a natural tendency to view younger students as suspect in their maturity and ability. Small differences in ability and performance could be more salient, and be magnified under this suspicion. Teachers might then be quicker and more willing to make referrals for retention or special education placements as a result. There is some evidence that this is true.

Gredler, (1983) noted that the literature of different countries often contains studies that bemoan the unreadiness of children who are the youngest in grade. This exists despite the fact that many have entry ages both younger and older than the American norm. He compares reading

achievement scores of English and American first graders and notes that the reading scores of the English children are higher, despite being a whole year younger.

Clearly, least some of the age effect can be attributed to the relative expectations of teachers, not to the performances of the students. That teachers' expectations are at work is also witnessed in the differential effects of age for males and females, and blacks and whites observed in nearly every study. Most authors point out that the differences result from the differential maturational rate of males and females. Yet in the Langer study, differences for males and females did not hold true for the black sample. While black males were retained at higher rates than black females in early grades, no predictive power accrued for achievement. Since retention is generally a teacher initiated action, this suggests that teacher attitudes and expectations differ for black males and females. Indeed, it seems likely that for males in general the teacher fully expects them to do less well than females. Behavioral differences and sex role stereotyping may be the root of such expectations, increasing the strength of the teacher's conception that young males (and particularly young black males) may experience difficulty adjusting to school.

A Preference for No Treatment: The Great Logical Leap

Despite the tenuous nature of the evidence, school readiness researchers plow on to the conclusion that for many students, especially males, a future of problems may be averted by holding the student out a year to gain maturity (e. g. Uphoff and Gilmore, 1985; Ilg and Ames, 1965; Donofrio, 1977). An objection to this conclusion can be framed to this assertion on two grounds.

The first is that such a recommendation implies that doing nothing is better than addressing the supposed problems. It is the only place in special education literature that we know of where no early intervention is presumed preferable to assessment and remediation. Driven to a logical conclusion, the advocates of such a course should posit that we need a differential age entry cutoff for blacks and whites, for the rich and poor, as well as for males and females. In fact since we know that retarded

children identified through early screening will experience lower achievement, and increased possibilities for retention and referral to special education, it might be just as wise to keep them out of school for a year (or two or three) in order to raise their maturity and ability to handle the stress of school experience. In fact, it is possible to determine from the data generated in these studies that rich white students, and students with a high IQ will benefit more from staying out of school a year than other groups do, because the initial benefits they have in predicted achievement will actually increase measured against younger groups (Green & Simmons, 1967).

Secondly, there exists evidence that merely holding students out of school does not increase their relative performance when they do enter. (DiPasquale et al., 1980) pointed out that students who were held out of school in his study received lower grades, were ranked lowest by peers on several sociometric categories, and were behind in achievement. DiPasquale et al., (1980) studied children screened and recommended for a year delay in school entry. They matched students whose parents had complied and students whose parents had enrolled students anyway with students in the normal range on the readiness instrument. Results showed that at risk children who entered school in spite of the warning trailed normal children in achievement. Yet, their results were significantly greater than those achieved by students held out a year at the same grade level testing. The differences were maintained over the 3 years covered by the study.

The authors of the studies are not unaware of the illogic of their positions. Langer et. al. (1984) for example, indicate that the absence of a treatment may have negative impact on some groups, notably blacks, and provide a recommendation different for black males who fail screening tests. Yet, these authors don't seem to recognize the irony of this position. If blacks are more at risk, the chop logic of these studies would dictate holding them out longer.

If a five year old handicapped child received this sort of advice, the parents might contemplate a law suit in subsequent years. The current thrust of research, practice, and legislation (e. g. Federal

Law 99-457) has been to provide service at appropriate ages for individually determined needs, regardless of minimal/maximal school attendance ages. We feel that the efforts to fix general and arbitrary age limits for school attendance is deleterious to students who are at risk, and for students of high ability.

Early Entrance and the Gifted: The Fallacy Continues

The fact that such logical errors have persisted, and indeed have been perpetuated in policy decisions in a large number of districts is dismaying. Some states and many districts have changed cutoff dates for entry, and many more routinely recommend "academic redshirting" for students deemed too immature to succeed (Frick, 1986). Yet the fallout doesn't stop here. Some districts are seriously questioning allowing gifted students to enter school early, or be accelerated through the grades once enrolled. Policy makers have seized on the notion that younger students in a grade have more difficulty in school and applied it to students who are well above their chronological age in ability. Misinterpretation and misapplication of poor literature has a negative impact on the quality of these students' education (cf. Maddux, 1983).

The evidence that younger children are more at risk for referral and labeling as LD is weak and causally tenuous. Even if the link were more securely established, it would argue for earlier, not later schooling. Yet, in their zeal to raise school entry dates, researchers have promoted the notion that even for gifted children, too early a school entry age bodes ill. These conclusions, however, are unwarranted on the basis of their own evidence, and are refuted by a large body of research literature in the field of gifted education.

In the first place, the studies in this field are particularly inappropriate for gifted students because, most often, they exclude the gifted from consideration. In those studies, for example, that look at referrals and retentions, the students referred for exceptional ability, or promoted ahead of chronological placement are removed from the analysis (Gredler, 1985). Usually, when ability is

considered at all, the analysis is ex post facto, averaging IQ across samples of the younger children in grade, rather than examining samples matched for IQ in various grade placements (Uphoff and Gilmore, 1985). The samples that do include gifted children generalize statements across the entire range, and do not take into account the highly selective nature of the process in which a child is considered for early admission or acceleration.

Secondly, these decision makers ignore a large body of research that finds early admission to be a valid programming option for students selected on the basis of academic achievement and maturity. Daurio (1977) reviewed literature on early admission and acceleration and found no evidence that students screened and given early admission experienced social or emotional harm, or that such students suffered declining performance in comparison with chronological or grade peers. Kulik and Kulik, (1983) conducted a meta-analysis of studies on relative academic achievement of accelerates and found that they maintained initial advantages over their chronological peers in every academic area and continued at the top range of achievement within their new grade placement.

Thirdly, the solutions adopted by states and districts to establish arbitrary, and invariant entry cutoffs ignore the variability of students at any age. Any effort, no matter how well intentioned and well designed, to establish some sort of optimum school entry age will be confounded by groups of students born before this date capable of high achievement in the curriculum. The demand for uniformity that public school education often requires tends to ignore the wide range of ability and developmental differences among the children that are their constituents.

Needs and Characteristics, Not Cutoffs and Birthdays

The problems of the gifted child faced with this type of policy enactment is actually reflective of the nature of this research in general. When faced with the apparent finding that an educational program does not meet the needs of some sub-group, the response is to determine a way to make the subgroup more like everyone else instead of addressing the educational characteristics of the group in question.

If young students are having difficulty, make them older. It does not occur to these people to examine the curriculum or the teachers to determine why these students are not meeting expectations. As

Gredler (1984, p.12) observed:

Carefully planning beginning instruction, individualized instruction with care and forethought is the primary need. School psychologists and others need to stop using age as an excuse as to why a child is not reading.

References

- Baer, C.J. (1956). The school progress and adjustment of underage and overage students. Elementary School Journal, 55, 331-336.
- Bigelow, E. B. (1934). School progress of underage children. Elementary School Journal, 35, 186-192.
- Carter, L. B. (1956). The effect of early school entrance on the scholastic achievement of elementary school children in the Austin Public Schools. Journal of Educational Research, 50, 91-103.
- Cook, T. D. & Campbell, D. T. (1979). Quasi-experimentation: Design and analysis in field settings. Chicago: Rand McNally.
- Daurio, S. P. (1979). Educational enrichment versus acceleration: A review of the literature. In W.C. George, S.J. Cohn, and J. Stanley (Eds), Educating the gifted: Acceleration and enrichment. (pp. 13-63), Baltimore: Johns Hopkins University Press.
- Department of Health and Human Services, Centers for Disease Control, Suicide Surveillance Summary: 1970-80. (March, 1985). Washington D.C.
- Diamond, G. H. (1983). The birthdate effect - A maturational effect? Journal of Learning Disabilities, 16, 161-164.
- Dietz, C. & Wilson, B. J. (1985). Beginning school age and academic achievement. Psychology in the Schools, 22, 93-94.
- DiPasquale, G. W., Moule, A. D., & Flewelling, R. W., (1980). The birthdate effect. Journal of Learning Disabilities, 13, 234-238.
- Donofrio, A. F. (1977). Grade repetition - Therapy of choice. Journal of Learning Disabilities, 10, 349-351.

- Forester, J. J. (1955). At what age should a child start school? School Executive, 74, 80-81.
- Frick, R. (1986). In support of academic redshirting, Young Children, 412, 9-10.
- Gerber, M. M. & Semmel, M. L. (1984). Teacher as imperfect test: Reconceptualizing the referral process. Educational Psychologist, 19, 137-148.
- Gredler, G. R. (1978). A look at some important factors in assessing readiness for school. Journal of Learning Disabilities, 11, 25-31.
- Gredler, G. R. (1980). The birthdate effect: Fact or Artifact? Journal of Learning Disabilities, 135, 9-12.
- Green, D. R., & Simmons, S. V. (1962). Chronological age and school entrance. Elementary School Journal, 63, 41-47.
- Hall, R. V. (1963). Does entrance age affect achievement? Elementary School Journal, 63, 391-396.
- Huff, S. (1984). The pre-kindergarten assessment. A predictor of success of early and late starters. Ed.S research project at Wright State University, Dayton, OH.
- Ilg, F. L. & Ames, L. G. (1965). School readiness: Behavior tests used at the Gesell Institute. New York: Harper & Row.
- King, I. B. (1955). Effect of age of entrance into grade 1 upon achievement in elementary school. Elementary School Journal, 55, 331-336.
- Kulik, J. A., & Kulik, C. C. (1984). Effects of accelerated instruction on students. Review of Educational Research, 54, 409-425.
- Langer, P., Kalk, J. H., & Searls, D. T. (1984). Age of admission and trends in achievement: A comparison of blacks and acucasians. American Educational Research Journal, 21, 61-78.
- Maddux, C. D. (1983). Early school entry for the gifted: New evidence and concerns. Roeper Review, 5(3), 15-17.
- Mawhinney, P. E. (1964). We gave up on early entrance. Michigan Education Journal.
- Miller, W. & Norris, R. C. (1967). Entrance age and school success. Journal of School Psychology, 6, 47-60.
- May, D. C., & Welch, E. L. (1984). The effects of developmental placement and early retention on children's later scores on standardized tests. UPsychology in the Schools, 21, 381-385.
- Partington, H. M. (1937). The relationship between first grade entrance age and success in the first three grades. National Elementary Principal, 16, 298-302.

- Pegnato, C. & Birch, J. (1959). Locating gifted children in junior high schools: A comparison of methods. Exceptional Children, 25, 300-304.
- Pollins, L. D. (1983). The effects of acceleration on the social and emotional development of gifted students. In C. P. Benlow and J. C. Stanley (Eds.), Academic percocity: Aspects of its development pp. 160-179). Baltimore: Johns Hopkins University Press.
- Robbins, T. (1984). Jitterbug Perfume, New York: Bantam Books.
- Southern, W.T. & Jones, E. D. (1987). Practitioner objections to the accademic acceleration of young gifted children. (Manuscript submitted for publication).
- Stufflebeam, D. L. & Webster, W. J. (1980). An analysis of alternative approaches to evaluation. Educational Evaluattion and Policy Analysis, 2, 5-20.
- Terman, L. M. (1925). Mental and physical traits of a thousand gifted children. In L. Terman (Ed.), Genetic studies of genius (Vol. 1). Stanford: Stanford University Press.
- Uphoff, J. K. & Gilmore, J. (1985). Pupil age and school entrance - How many are ready for success? Educational Leadership, 43(1), 86-90.
- Uphoff, J. K. & Gilmore, J. (1986). Pupil age and school entrance - How many are ready for success? Young Children, 41(2), 11-16.
- Uphoff, J. K. & Gilmore, J. (1987). Pupil age and school entrance - How many are ready for success? Dimensions, 8(7), 3-5.

Tables I and II from: Langer, P., Kalk, J. H., & Searls, D. T. (1984). Age of admission and trends in achievement: A comparison of blacks and acucasians. American Educational Research Journal, 21, 61-78.

Tables I and II from: Langer, P., Kalk, J. H., & Searls, D. T. (1984). Age of admission and trends in achievement: A comparison of blacks and acucasians. American Educational Research Journal, 21, 61-78.

TABLE I
Summary of Multiple Regression for Age 9 Caucasians Combined for the Mathematics, Science, and Reading Assessments

Variable	B	F	p	R ²	Adj. R ²	Simple r	Overall F	p
Relative age	-.015	30.9	.01	.002		-.048	20.4	.01
Class age	.069	51.6	.01	.003	.003	.013	20.4	.01
Home environment low	-.264	163.6	.01	.047	.047	-.211	230.0	.01
Parental education low	-.242	137.7	.01	.077	.077	-.211	291.2	.01
Type of community high	.237	90.5	.01	.089	.088	.122	269.9	.01
Home environment high	.206	120.0	.01	.097	.097	.201	249.5	.01
Southeast	-.159	59.2	.01	.103	.102	-.094	227.1	.01
Parental education high	.156	56.1	.01	.106	.106	.199	206.9	.01
Type of community low	-.118	23.0	.01	.108	.107	-.075	187.0	.01
Northeast	.086	16.7	.01	.109	.108	.073	170.1	.01
Sex	-.004	.2	ns	.109	.108	.017	154.7	.01

TABLE II
Summary of Multiple Regression for Age 9 Blacks Combined for the Mathematics, Science, and Reading Assessments

Variable	B	F	p	R ²	Adj. R ²	Simple r	Overall F	p
Relative age	-.017	9.4	.01	.003		-.059	6.0	.01
Class age	.056	7.6	.01	.004	.003	.006	6.0	.01
Home environment low	-.184	20.6	.01	.032	.031	-.171	33.9	.01
Type of community high	.441	27.6	.01	.050	.048	.142	39.7	.01
Parental education low	-.160	13.9	.01	.062	.061	-.144	40.4	.01
Southeast	-.148	13.2	.01	.069	.068	-.093	37.8	.01
Type of community low	-.183	22.1	.01	.075	.073	-.100	35.4	.01
Home environment high	.177	13.5	.01	.080	.077	.157	33.0	.01
Northeast	.159	9.5	.01	.083	.080	.096	30.5	.01
Sex	-.045	6.6	.01	.085	.082	-.034	28.1	.01
Parental education high	.120	6.4	.01	.087	.083	.133	26.2	.01